

# JAMS: A JSON Annotated Music Specification for Reproducible MIR Research



Eric J. Humphrey<sup>1</sup>, Justin Salamon<sup>1,2</sup>, Oriol Nieto<sup>1</sup>, Jon Forsyth<sup>1</sup>, Rachel M. Bittner<sup>1</sup> and Juan Pablo Bello<sup>1</sup>  
<sup>1</sup>Music and Audio Research Laboratory, New York University {ejhumphrey, justin.salamon, oriol, jpf211, rachel.bittner, jpbello}@nyu.edu  
<sup>2</sup>Center for Urban Science and Progress, New York University

## Context

Music annotations are a critical component in a variety of active research areas in MIR.

- ▶ Used for both development and evaluation
- ▶ Time scales at the level of a recording, e.g. beats, chords, tags
- ▶ Typically, “lab” files or other custom plain-text conventions are used to serialize this data

## So, what’s the problem?

Modern research trends are demanding more of music annotations.

- ▶ Increasingly complex annotations
- ▶ Multiple annotations per task
- ▶ Multiple annotations per recording

Importantly: This data is structured.

## Key Design Criteria

### Simplicity

- ▶ Easy to use, easy to understand

### Structure

- ▶ Mixed data types, e.g. scalars, strings, arrays of different shapes

### Sustainability

- ▶ Methodology and metadata should be embedded in the annotation itself
- ▶ Leverage standardized tools and technologies developed by larger communities

## Introducing JAMS

A **JAMS Object** contains:

- ▶ arrays of **Annotations** for different tasks, e.g. beat or chord estimation
- ▶ a file metadata object
- ▶ a global sandbox

An **Annotation** contains:

- ▶ a data array corresponding to the task, e.g. beat times or chord intervals
- ▶ an annotation metadata object, specific to this particular annotation
- ▶ an annotation sandbox

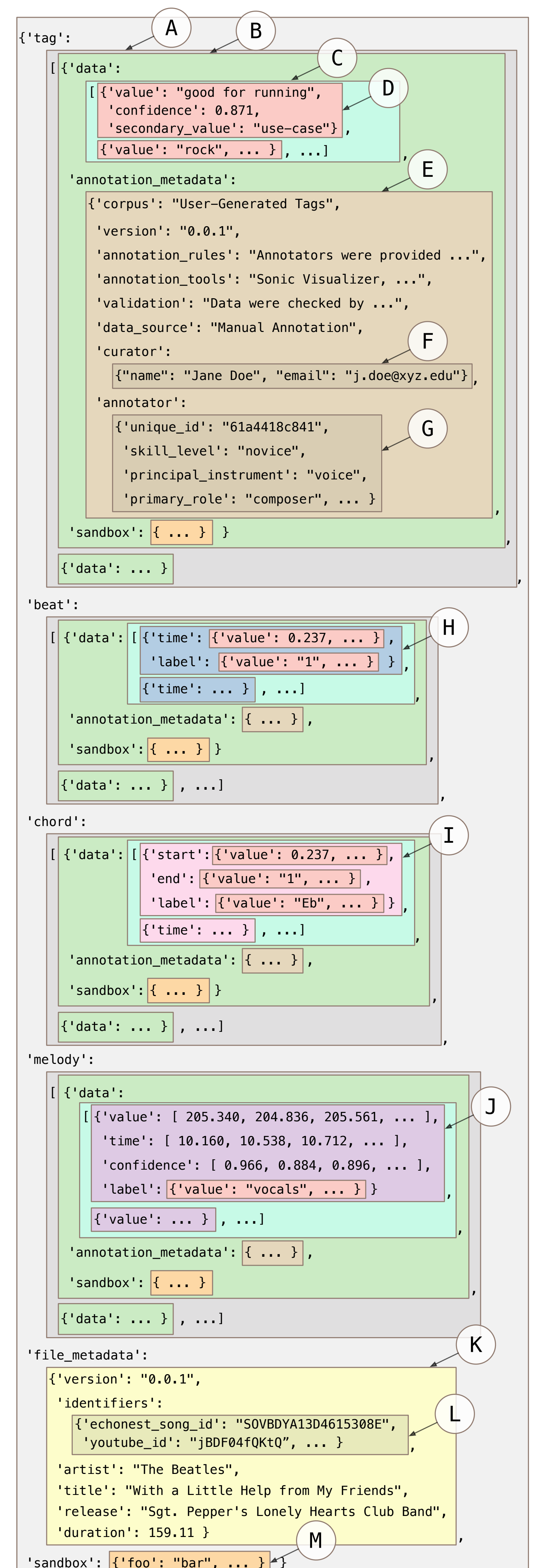
Data arrays contain a collection of similar types:

- ▶ **Observations** are global descriptors, e.g. tags or genre
- ▶ **Events** are instantaneous occurrences, e.g. beats or onsets
- ▶ **Ranges** describe intervals in time, e.g. chords or segments
- ▶ **TimeSeries** represent dense data, e.g. melody

## An Illustrated Example

Unique JAMS Objects

- A. Annotation Array
- B. Annotation
- C. Data Array
- D. Observation
- E. Annotation Metadata
- F. Curator
- G. Annotator Sandbox
- H. Event
- I. Range
- J. TimeSeries
- K. File Metadata
- L. Identifiers Sandbox
- M. Global Sandbox



## Practical Considerations

Software APIs

- ▶ Python and Matlab
- ▶ Baseline parsers to convert data

Pre-converted Datasets

- ▶ Isophonics (beat, chord, key, segment)
- ▶ Billboard, tmc323 (chord)
- ▶ SALAMI (two-layer segments)
- ▶ RockCorpus (chord, key, measures, melody)
- ▶ Cal500 / 10k (tag)
- ▶ ADC04, MIREX05 (melody)
- ▶ MedleyDB (melody, pitch, source)

## Next Steps

Calling all users, contributors and collaborators!

- ▶ Make this a community endeavor
- ▶ Actively seek out and incorporate feedback

Limitations and Near-field Challenges

- ▶ Refine datatypes
- ▶ Tighter JSON support in Matlab
- ▶ Scalability and streaming JSON

“Ground Truth” is tenuous in MIR

- ▶ Embrace multiple reference annotations
- ▶ Subjectivity is variable; all perspectives are valid

Future goals

- ▶ Integration with other tools, e.g. Sonic Visualizer, librosa, mir\_eval, and MIREX
- ▶ Better annotation coverage for existing datasets
- ▶ Establish a freebase-style archive of annotations